Estimation and Use of Load Threshold for Energy Saving in Wireless Networks

Harri Saarnisaari
Centre for Wireless Communications
University of Oulu
Oulu, Finland

Abstract—**Energy saving is an increasing trend in wireless communications development activities. This paper investigates how historic traffic load data can be used to estimate and set the load threshold used for switching on or off base stations or their features on a particular weekday. The use is enhanced by an abnormal day detection feature that is used to prevent undesired power on/off actions on days were users do not follow their usual behavior. Power on and off are also prohibited due to impulsive, i.e., single, observations that exceed state change thresholds. The simulation results show that the method estimated power on/off moments according to desired load level very well, especially if the load variation is limited, while preventing undesired power on/off actions.**

I. INTRODUCTION

Energy saving (ES) is an emerging trend in wireless communication system development already resulted a rather rich literature, e.g., [1], [2], [3], [4], standardization activities, e.g., [5] and it is included into next generation system visions [6]. The reasons for ES are well described in the referenced documents and, as a consequence, not repeated herein. An efficient ES approach is to switch off unnecessary access points (APs) but also switch them on once needed. One possibility is to allow micro, i.e., very short term, sleeps investigated, e.g., in [1]. Naturally, micro sleeps concerns active usage periods and care has to be taken to avoid ping pong between on and off [7]. Very often base coverage is guaranteed by a macro cell and additional features either to that AP or additional APs are installed for supporting higher load requirements like hot spots in the area. Therefore, in a big picture, switching on and off APs and their features for active periods and low load periods, respectively, is of interest.

Standardization bodies have concluded, so far, that there could be operations, administration and maintenance (OAM) or signalling based process for ES. All these could use load thresholds to decide activation or deactivation of ES functionalities and perform more precise adjusting (e.g., order to switch on or off) [8]. The determination or use of the load threshold is not discussed, however, but left for vendors or operators.

The determination of the load threshold using probability maximization of ES is discussed in [9] but, as shown in that paper in its further studies, application of it to more practical systems needed ad hoc adjustment of the threshold. Some switch on or off strategies have been discussed in the literature. [7] discusses a distributed mechanism where APs are switched on or off one by one after negotiation between the neighbors. However, any threshold setting strategy was not discussed though impacts of the allowed maximum load on APs were shown. A one by one switch on or off strategy is also discussed in [10]. These allow also short term sleeps and may result the ping pong phenomena. If the big picture is the desired aim then the recognition or prediction of initial load growth or final diminution should be pursued. A short term load prediction based approach is proposed in [11].

This paper contributes by providing a robust estimation of the threshold based on historic load data and a novel historic load based approach for the threshold setting in the big picture case. In addition, discussions to its adaptation based on operation time characteristics are provided. After introducing the a novel switch on or off strategy simulations are used to show its performance in some cases, e.g., when the actual load level varies significantly.

II. ENVIRONMENT DESCRIPTION, SYSTEM REQUIREMENTS

Assume an area that is served by an access point (AP) or few APs at low load situations. This set of APs and their features is called the basement level or level 0. Once traffic load increases new features of current active APs and/or new APs are activated. First, just a subset of capabilities may be called and these form, together with the basement level, level 1. This could continue gradually until all resources are used, or all L levels $0, \ldots, L-1$. Once the load is decreasing APs are shut down. Finally, only the basement level remains. A question is, naturally, that when switching on and off processes should be initiated.

Obviously, on the rising load case, the first action is to switch APs and/or their features on after which the existing users could be reassigned to an appropriate APs according the usual (SON) mechanisms and new users could be served better without QoS reduction. On the decreasing load case the users must first be removed to still remaining APs and the proper shut down process can be started after the AP is empty. Therefore, the rising and falling load cases are a bit different. In the latter reactions related to users must be started and finalized before the shut down whereas in the former they can be started after powering the devices on. As a consequence, the

switch off process includes emptying the APs (to be switched off) and shut down them. The switch on process includes power on the APs (on the next level) since reassigning the users to different APs is normal operation.

As base stations are assembled into an area one of the parameters that will be set, in the future, is the load threshold for energy saving (ES) functionalities that is in part responsible on the switching on or off [8]. Initially this could be based on hands-on experience but once operation is running more precise values should be derived. This could be based on the historic data on the area, which is the aim in this paper.

A. Other Considerations

It is perfectly understood that the load threshold is not the sole parameter that (de)activates ES actions or that switching on and off is the only needed action [8]. Furthermore, there are other use possibilities and scenarios too. Some considerations around the topic are discussed herein.

Obviously, the load measurements and the load threshold setting could be AP or even AP sector based, or they could be based on broader, area based analysis and setting. Similarly, the control could be centralized (OAM based) or distributed to APs.

The learning of usual loads could be used not only for the load threshold setting and adjusting but also for deciding the number of layers and the layer contents, i.e., for the resource management (SON) functionalities. They could also be used for (fine) tuning the thresholds, including the critical threshold, based on QoS or other measurement from the network.

As already discussed, at the final stage, one has to decide are the micro sleeps inside the levels allowed or not. In addition, one has to decide is the level (gradual) based approach discussed herein or the continuous, AP by AP, approach more appropriate. The latter allows continuous micro sleeps but may result a ping pong effect.

III. ESTIMATION OF LOAD THRESHOLD

User density and activity depend on the time of the day, weekday, month and may change in an area over the years. Thus, the estimator must be adaptive. Since learning based on monthly features takes years, weekly based system is proposed and public holidays, vacations, strike days etc are handled as deviancy in the estimation but should be treated so also when used. Weeks long estimation periods are acceptable since the operation time period of APs is typically years such that improved operation could be maintained most of the life span.

Therefore, time is divided into intervals, e.g., 15 min intervals, per day and prediction will be based on several weeks of observations. Initially, the (adaptation) process may be speeded up assuming that working days are similar. In some areas (e.g., office areas) weekend days could be jointly similar whereas in some they differs, e.g., if Saturday is a shopping date but Sunday is not. In those cases it might be better to wait until daily estimate is ready and use the initial values before that.

It should be ensured that the estimation process is initially not including vacation or holiday periods since those results false conclusions. Later on, once estimator is ready, abnormal measurement results could be ignored. Of course, if known, typical vacation and public holidays can be feed into the system and used to aid avoiding abnormal days in updating the estimation result.

Since there is natural variation in load between two consecutive same weekday, e.g., Monday, it might be necessary to estimate both the mean and variance of the load.

Basically, any suitable estimation algorithm of the mean will do the job. The principles are described herein.

1) Collect a basic sample set for that particular day for all intervals, N_w samples for each interval.
2) Check does there exist outliers and remove them (robustness).
3) Compute the mean and variance. These are the initial estimates.
4) As new data sample arrives, check that it is not an outlier and, if not, update the estimates using a forgetting (moving average) estimator. If it is an outlier, do not update. Update could be $x_i = \lambda x_{i-1} + (1 - \lambda)m_i$, where x_i is the estimate, m_i the measurement and λ the forgetting factor.

Outlier detection algorithms have been considered, e.g., in [12]. A question is how many (N_w) samples are needed and another what should the forgetting factor λ be. Probably forgetting could be slow. If sudden permanent changes occur those could be detected counting the number of consecutive outliers for that moment of the day. If that number exceeds typical vacation period, or do not occur during typical vacation period, then new estimates should be computed since for some reason user behavior on the area has been changed, e.g., a big office has been closed or opened. For fast update a second memory may be used that computes the average based on those outliers or based on all samples where outliers have not been removed.

With 15 min intervals there will be a $24 \times 4 \times 7$ matrix that is to be updated. If longer intervals could be used the storage space needed for the estimator would be smaller. If backup computations described in the previous paragraph are used two matrices are needed.

IV. USE OF LOAD THRESHOLD

Once the load is estimated daily basis it could be used to steer the threshold. In what follows, a few possibilities are considered.

A. Basic Setting Rules

At the basement level, and other levels too, it is known after which load, a critical load level, quality-of-service (QoS) will be decreased or connections will be dropped etc. Therefore, the main goal is to set load thresholds T_L such that switch on and off processes could start on time. At the rising edge switch on should be ready before the critical load is achieved. Assume that the switch on duration (such that APs are ready to serve users) is D_{on} seconds. The rising edge threshold $T_{L,re}$ is the estimated load D_{on} seconds before the critical load is usually achieved on that particular weekday. Also the moment on that particular weekday when the load is crossed could be used
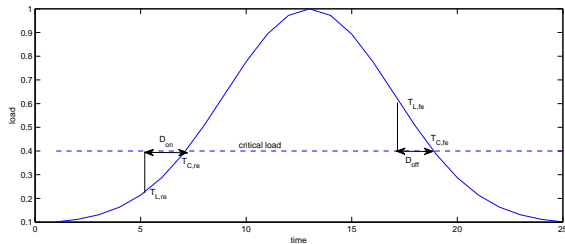
Fig. 1.   The illustration of the proposed threshold setting process.

as an indicator to start the power on process. At the falling edge, the shut down can be started when the load is under the critical load, which means that only preparatory actions need to be counted on duration D_{off} when setting the threshold $T_{L,fe}$ or the corresponding moment. In real life, in the switch off, since removing of users to remaining APs may not be ready (depending on traffic change rate) at the point when the critical load is achieved handovers may take certain time after that too, i.e., the actual switch off occurs later.

One must also remember that there are variations on user behavior and timing such that it might be a good practice to use the worst case (e.g., mean$\pm2\times$standard deviation) as a guideline instead of the mean. It appears that it is the $+$ sign since at the rising edge the maximum possible load achieves the critical load first and at the falling edge the maximum possible load achieves the critical load last.

B. Norm Day

The above described load threshold or time-of-the-day setting should work on normal particular weekdays, so called norm days. However, all days are not norm days since there may be strikes, unexpected public holidays, polling days, some popular events at weekends on otherwise almost quiet spots, etc. On those abnormal days the set (norm day) thresholds may not work.

For those cases other load thresholds, say T_C, are needed, probably closer to the critical load, even equal to those, and once these are crossed the switch on or off processes are started. Fig. 1 illustrates the thresholds relative to the critical load.

One could naturally use these closer (backup) thresholds for usual operation but since they are closer the critical load than the user behavior based optimal thresholds some performance losses may occur. Indeed, the measurement based thresholds are more sensitive, i.e., their usage makes the system more sensitive based on observations. In this sense they adopt the brain model by [13]. Therein, it is assumed that if brains have made an observation that likely has certain consequences, they make observation of that consequence more easily, e.g., the brains have seen writing LITTL such that the threshold for observing the letter E will be decreased, i.e., the probability of detecting the letter E will be increased.

How to use the norm day concept? If on that particular day the norm day threshold is crossed at the right moment it is decided that it is the norm day and that threshold initiates

the switch process. Alternatively, it may occur that the norm day threshold is crossed early, later or not at all. These are considered as abnormal days and the thresholds T_C are used. In practice this means that it must be checked that crossing really occurs on norm days. At the rising load case, since the power on should be initiated early this means that the power on process will be terminated if the day appears to be an abnormal day. Naturally, this power on and its termination wastes some energy but since the next level should be ready before the critical load is achieved it might be a necessary step.

a) Sudden Change: If the change is sudden then suitable load threshold cannot be found but instead a stable load value (from a plateau). In this case a shorter time interval may help. For example, adaptive measurement duration where duration is made shorter on load change times. Or threshold based on time of the day, i.e., it is known that typically the switch process should be started at X o'clock. This might be problematic if the day appears to be an abnormal day. In that case the switch on or off process will be initiated although there is no need for that. Indeed, this is worse in the switch on since the APs and features are powered on.

V. PERFORMANCE

A. Estimator

An important question is how many weeks are needed for good estimates. Let $m(t_i)$ denote the random variable describing the measurement in an interval t_i. It is assumed to be a Gaussian variable with mean $M(t_i)$ and variance $\sigma^2(t_i)$. Their estimation is a standard estimation problem and the behavior of the estimator is well known. The Cramér-Rao bounds for the mean and variance are inversely proportional to the number of samples N [14]. Indeed, the so called two-sigma accuracy, which means that 95 % of the measurements are within these values, for the mean is $M(t_i) \pm 2\sigma(t_i)/\sqrt{N}$ and for the variance it is $\sigma^2(t_i) \pm 2\sigma^2(t_i)\sqrt{2/N}$. Therefore, the variance is harder to estimate but if it is small related to the mean its effect is not so significant.

If the starting question in this paragraph is considered, it is seen that two months (about 9 weeks) observation period gives 3 fold improvement compared to a single shot case. For example, assuming that the standard deviation is 20 % of the mean it follows that the previous accuracy expressions become, after 9 weeks, as $M(t_i) \pm 0.13M(t_i)$ and $0.04M^2(t_i) \pm 0.05M^2(t_i)$, respectively. If the latter is transformed to the standard deviation its maximum value is $0.3M(t_i)$. Early it was proposed that $M(t_i) + 2\sigma^2(t_I)$ is used in the threshold setting since it provides the worst case from the load's perspective. Therefore, including errors in the mean and variance, the total maximum error in this case is highly probable to be within $0.4M(t_i)$ that is the original 2σ value in this case. As a conclusion, if the load variation is modest the few months initial estimation period should be adequate.

B. Threshold Setting

Since this is initial work on the subject and since the real working procedure could be made very complicated the simulations are restricted. A simple threshold adjustment algorithm,
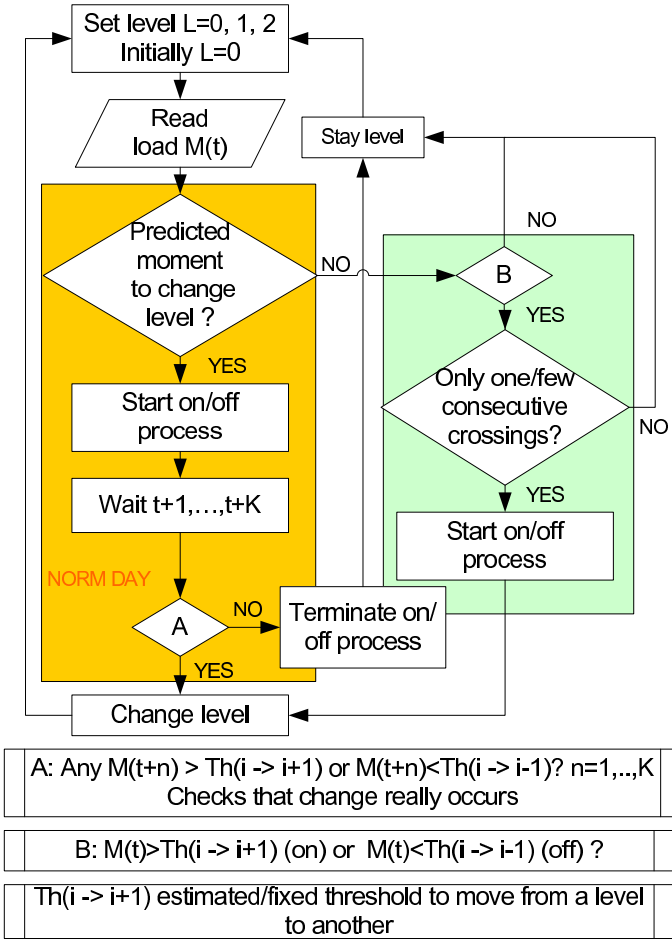
Fig. 2. The flow chart of the level change algorithm.

Fig. 3. The mean of the norm weekday and abnormal weekend day. The normal weekend days have flat 10 % load.

whose flow diagram is shown in Fig. 2, is used instead of fully developed one. The algorithm assumes a norm day but checks does the measured load exceed the threshold in the next few samples (now one) before it decides to change the level. At the rising edge this means terminating the power on process that was initiated according the norm day threshold. Indeed, at a norm day the process will be initiated at the estimated moment at the rising edge and one sample after the estimate moment at the falling edge. The estimator predicts the last moment before the crossing at the rising edge and the first moment after the crossing at the falling edge. At moments not predicted by the norm day estimator, i.e., at all other moments, the load thresholds 15 % and 55 % are used to initiate the power on or off processes.

The simulation covers 12 weeks divided into 15 min intervals. All working days have the same profile as well are weekend days equal. The profiles show the mean but generated load could vary 0 % and 20 % from the mean according to Gaussian distribution such that the variation is the standard deviation. This means that at high load the variation is higher. However, values are kept within 0 % to 100 % $[0 \ldots 1]$. Abnormal weekdays are like normal weekend days (where load is steady 10 %) whereas at abnormal weekend days have an few hours lasting increase. The means of normal weekday
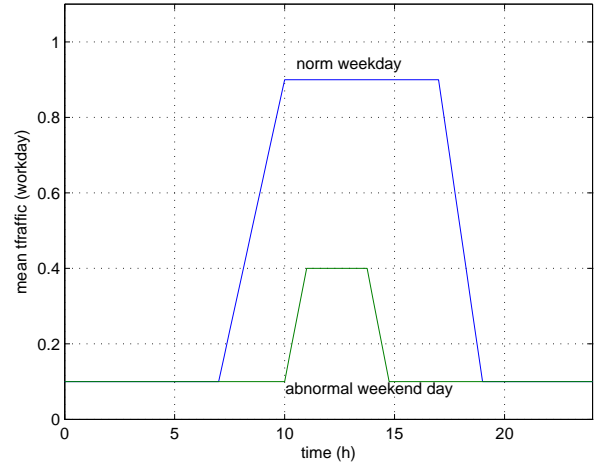
and abnormal weekend day are shown in Fig 3.

In addition, the following constraints are used in these simulations: i) The training period of 4 weeks does not contain abnormal days. The estimator's forgetting factor is 0.97 since longer observation periods improve the estimator. Outliers are not detected but estimates are not updated on abnormal days. ii) At the basement level load is 0 - 15 %, the first level 15 - 55 % and the second level 55-100 %. When moving up, the algorithm predicts when to move based on the norm day, if observed, or the upper limit as a threshold otherwise. The going down algorithm waits or predicts when the load is below the critical load. iv) Abnormal days occur at week 12 at Monday, Tuesday, Saturday and Sunday. The estimator should not update it at all time intervals at abnormal days and the norm day detector should not react.

The results show the level setting results from the last two weeks showing one normal week and the abnormal week with different variation. The (red) squares present moments when the algorithm is on the basement level, black stars the level 1 and magenta diamonds the level 2. The two marks in a moment express the level change moment. The zero variation results in Fig. 4 show that the estimated moments are precise, as expected, i.e., one moment before the level at the rising edge and one moment after the level at the falling edge. Furthermore, the algorithm could predict correctly the abnormal days when the threshold is not crossed and uses the 15 % level based threshold at the abnormal days when the threshold is crossed, as it should.

The 20 % variation results in Fig. 5 show that at the rising edge the algorithm may make its decision too late, shown by arrows, at the norm days but, on the other hand, it does not care about occasional level crossings seen at all levels or, in other words, do not ping pong between levels.

VI. Conclusions

This paper shows that the estimated load threshold can quite reliably predict power on/off moments on so called norm days.
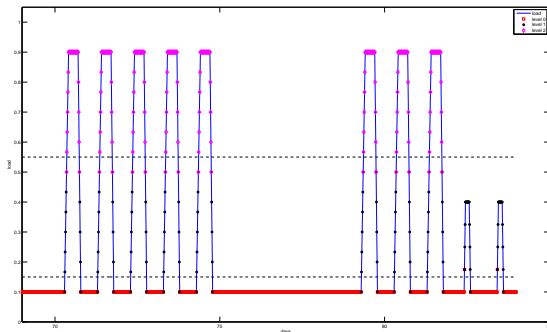
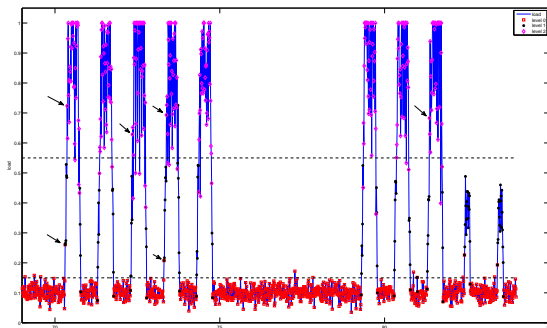Fig. 4. The selected levels with 0 % variation from the last two weeks.

Fig. 5. The selected levels with 20 % variation from the last two weeks. Arrows point to the problematic points where late decisions have been made.

The algorithm also shows that using a few consecutive samples to make the final level change decision helps to avoid ping pong between the levels, which was desired effect in this paper that considers the big picture. Some may prefer micro sleeps between the big moments where load changes for longer periods and then papers discussed in the introduction may be helpful starting points. Therefore, the idea presented in this paper could be a good starting point for algorithms needed in energy saving SON functionalities.

The problematic point seems to be the rising load case where natural load variation may cause late power on decisions using the algorithm. Therefore, work is needed to find a procedure to better predict or calculate how much earlier the predicted moment the power on process should start. The use of estimated standard deviation as discussed in the paper was trialled herein, but it caused problems at the basement level since the used variation was so large that the estimated threshold was often higher than the lower critical load level. Consequently, those results were not shown. In order to tune the algorithm to match better the practice real life load measurements could be helpful since they would show the practical variation and load profiles better than artificial examples used herein.

The purpose was to define the level change thresholds. However, these could be used for several other managements too. For example, the estimated and predicted load and its type

(voice, short messages, data) can be used to define the number, type and capacity of active access points inside the levels but that possibility was not elaborated in this paper. This means that the moments when to power on/off additional features for base stations, WIFI access (hot spot) points, etc. could be estimated better. Therefore, well predicted power on/off moments in addition to estimated and predicted load and its type open possibilities for several other energy saving as well as other management options.

REFERENCES

[1] J. Xu and C. Yu, "Improving network energy efficiency through cooperative idling in the multi-cell systems," *EURASIP Journal on Wireless Communications and Networking*, no. 165, 2011.
[2] O. Blume, H. Eckhardt, S. Klein, E. Kuehn, and W. Wajda, "Energy savings in mobile networks based on adaptation to traffic statistics," *Bell Labs Technical Journal*, no. 2, pp. 77–94, 2010.
[3] R. Li, Z. Zhao, X. Zhou, and H. Zhang, "Energy savings scheme in radio access netwroks via compressive sensing-based traffic load prediction," *Transactions on Emerging Telecommunications Technologies*, 2012.
[4] J. Lorincz, A. Capone, and D. Begusic, "Impact of service rates and base station switching granularity on energy consumption of cellular networks," *EURASIP Journal on Wireless Communications and Networking*, no. 342, 2012.
[5] "Potential solutions for energy saving for E-UTRAN (3GPP TR 36.927 version 11.0.0 Release 11)," ETSI: LTE; Evolved Universal Terrestrial Radio Acceess (E-UTRA), Tech. Rep. ETSI TR 136 927 V11.0.0 (2012-10), 2012.
[6] "Looking ahead to 5G: Building a virtual zero latency gigabit experience," Nokia Solutions and Networks, FutureWorks, NSN White Paper, December 2013.
[7] E. Oh, K. Son, and B. Krishnamachari, "Dynamic base station switching-on/off strategires for green cellular networks," *IEEE Transactions on Wireless Communications*, vol. 12.
[8] "Study on operations, administration and maintaenance (OAM) aspects of inter-radio-access-technology (RAT) energy saving (Release 11)," 3rd Generation Partnership Project, Technical Specification Group Services and System Aspects, Tech. Rep. 3GPP TR 32.834 V11.0.0 (2012-01), 2012.
[9] S. Wang and W. Guo, "Energy and cost implications of a traffic aware and QoS constrained sleep mode mechanism," *IET Communications*, vol. 7.
[10] E. Kisielius, A. Avramona, A. Zakrzewska, and S. Ruepp, "Energy efficiency in self organizing networks," in *OPNETWORK*, 2013.
[11] P. Kulkarni, T. Lewis, and Z. Fan, "Simple traffic prediction mechanism and its applications in wireless networks," *Wireless Personal Communications*, vol. 59, pp. 261–274, 2011.
[12] H. Saarnisaari, P. Henttu, and M. Juntti, "Iterative multidimensional impulse detectors based on the classical diagnostic method," *IEEE Transactions on Communications*, vol. 53, no. 3, pp. 395–398, March 2005.
[13] R. Kurzweil, *How to Create a Mind: The Secret of human thought Revealed*. Penguin Books, 2012.
[14] S. M. Kay, *Fundamendals of Statistical Signal Prosessing: Estimation Theory*. Englewood Cliffs, New Jersey: PTR Prentice Hall, 1993.